

VOLUME 64
NUMBER 5

WHOLE No. 311
1950

Psychological Monographs:

General and Applied

Combining the *Applied Psychology Monographs* and the *Archives of Psychology*
with the *Psychological Monographs*

HERBERT S. CONRAD, *Editor*

Changes in Responses to the Minnesota Multiphasic Inventory Following Certain Therapies

By

WILLIAM SCHOFIELD

*Departments of Psychiatry and Psychology
University of Minnesota*

Accepted for publication, October 7, 1949

Price \$1.00

Published by

THE AMERICAN PSYCHOLOGICAL ASSOCIATION
1515 MASSACHUSETTS AVE. N.W., WASHINGTON 5, D.C.

MAILED JAN 23
1950

MAILED
JAN 23

Psychological Monographs:

General and Applied

Emphasizing the Practical Application of Psychological Principles and the Solution of Psychological Problems

HERBERT S. CONRAD, Editor

COPYRIGHT, 1950, BY THE

AMERICAN PSYCHOLOGICAL ASSOCIATION

Changes in Responses to the
Minnesota Multiphasic Inventory
Following Certain Therapies

WILLIAM SCHORRILL

Department of Psychology and Psychiatry
University of Minnesota

THE AMERICAN PSYCHOLOGICAL ASSOCIATION

1201 SEVENTEENTH STREET, N.W., WASHINGTON 5, D.C.

ACKNOWLEDGMENT

THE WRITER gratefully acknowledges the encouragement and critical suggestions of Professor Starke R. Hathaway.

He is especially indebted to his wife, Geraldine B. Schofield, for many hours of assistance in the tedious task of recording the basic data.

The writer also wishes to thank the staff of the University of Minnesota Hospitals, Psychopathic Unit, for generous cooperation with the problems of collecting pre- and post-therapy data.

WILLIAM SCHOFIELD

TABLE OF CONTENTS

ACKNOWLEDGMENT	iii
I. INTRODUCTION	1
II. PURPOSE, SUBJECTS AND PROCEDURES	5
III. RESULTS	9
A. Mean Profiles for the Normal and Clinical Groups	9
B. Group Differences in the Frequency and Nature of Response Changes	15
C. Item Change in Relation to Score Change	19
D. Characteristics of the Most Stable Items and the Most Variable Items in Each Group	21
E. Investigation of a Therapy-Susceptible Scale	24
IV. SUMMARY AND CONCLUSIONS	28
APPENDIX	32
BIBLIOGRAPHY	33

TABLE OF CONTENTS

1. Introduction	1
2. The Problem	2
3. The Method	3
4. The Results	4
5. The Discussion	5
6. The Conclusion	6
7. The Acknowledgments	7
8. The References	8
9. The Appendix	9
10. The Index	10

CHANGES IN RESPONSES TO THE MINNESOTA MULTIPHASIC INVENTORY FOLLOWING CERTAIN THERAPIES¹

CHAPTER I

INTRODUCTION

THREE INFLUENCES stimulated the present study: (a) the intensive demand for more and better psychiatric services, (b) the attempts to formulate new systems of psychotherapy, and (c) the very recent but increasing efforts to establish objective criteria for the evaluation of therapeutic endeavors. Specifically, this study is concerned with an exploration of certain of the theoretical and practical problems involved in the measurement by means of a structured personality inventory of therapeutically induced changes in personality.

Psychiatry is currently experiencing a period of intense activity which is in direct response to the demands for therapy of the numberless personality disorders of all degrees engendered in civilian and veteran populations alike by the emotional stresses of World War II (1). With the addition of the personal maladjustments originating under the pressure of the emergency years to the pre-war case load, it is natural that the interest and contributions of the government and the public should be in the direction of increasing the number and quality of centers for the training of more clinicians. The approximately 4000 psychiatrists and 800 clinical psychologists now listed on the rolls of their respective organizations cannot, in point of numbers, handle in anything approaching adequate fashion the prophylactic and therapeutic needs of the population. Estimates have placed the number of psychiatrists required in certain areas at from four to seven times the number now available (5). It was primarily to remedy the dearth of adequately trained personnel that the National Mental Health Act was passed in 1946. Under the provisions of this act, the United States Public Health Service is directing and subsidizing

an extensive program to improve training facilities in psychiatry, clinical psychology, and psychiatric social work and to increase the number of trainees in these fields.

It is necessary, though difficult, in the midst of practical considerations and the exigencies of daily service to attempt a long view. Some workers already have expressed doubt as to the ultimate validity of the present emphasis on training. Kubie suggests that there is a fallacy in the theory that the training of more workers will suffice to improve the psychiatric health of the community or nation, at least as measured in terms of the number of patients (12). He points out that increased availability of services invariably leads to increased demand, so that the "catching up" point is never reached. An obviously crucial factor in determining the patients-per-therapist ratio is the effectiveness of the remedies available to the clinician. Not only must more workers be trained, but their training and therapeutic armamentarium must be enhanced by new techniques of proven efficacy. The development and testing of new techniques of therapy demand a research orientation, and research programs must not be grossly subordinated to the construction of curricula if psychiatry is to derive maximum benefit from the advantage it now enjoys under the triple incentives of social need, favorable public opinion, and government funds.

It is fortunate that the effective relationships which were established between psychiatry and clinical psychology in the armed services show promise of being maintained. Clinical psychology also is benefiting from government subsidization and while this support is, again, primarily to increase the number of trainees in this field, it may be hoped that a direct impetus to research will result.

It is only through the medium of research, supported and conducted by all professional workers concerned with the treatment and prevention of psychiatric disorders, that we may hope to achieve the construction and sharpening of those diagnostic, prognostic, and evaluative instruments which are necessary to assure the most efficient application of such facilities as are available for the care of personality disturbances at all levels of their manifestations.

¹This paper is a part of a thesis submitted to the Department of Psychology of the University of Minnesota in partial fulfillment of the requirements for the degree of Doctor of Philosophy, June, 1948.

What are the most pressing research needs in the field of mental health from a long-term view? First of all, there is a need for research which will establish in a definitive manner the level of effectiveness of existent therapies with respect to specifically diagnosed disorders. Secondly, there is a need for research which will establish the validity of newly proposed therapies. Inherent in researches of this type is the specific problem of determining which patients improve, and how much they improve, and which patients do not improve under a given regimen. The solution of this problem is contingent upon the attainment of a standard, objective, representative evaluation of the status of an individual's personality.

This paper presents an investigation of some of the problems arising in the application and interpretation of such evaluations, particularly as they furnish data on the effectiveness of given therapies.

The histories of the major psychiatric therapies have followed a certain pattern. In brief, there has been an initial enthusiastic announcement of the new technique accompanied with or followed shortly by figures on recovery which have represented not merely a sizeable improvement over those for earlier therapies but which have approached the 100 per cent point so closely as to suggest a specific cure. Later reports on the technique from other clinics and sometimes from the clinic of origin have shown considerably lower frequencies of recovery. Thus, Sakel in first publishing on insulin shock therapy of schizophrenia claimed 87.9 per cent remissions. This figure has not been duplicated by any other workers.

There are undoubtedly many factors involved in the explanation of diver-

gences between comparable figures for recovery rates in two clinics as well as for the failure of originally very high rates of remission in a given clinic to be maintained in further studies. A considerable part of divergence in interclinic statistics is probably due to differences in the systems (or lack of system) by which the status of patients is evaluated both before and after a course of therapy. Most commonly such evaluations are characterized by subjectivity, with little concern that the evaluation of the patient before and after therapy be on the same schedule of variables (3). In extreme form, this type of evaluation of a patient's progress is expressed in terms of the facts that his diagnosis at time of admission to the hospital was "schizophrenia—paranoid type" and that he was noted at time of discharge to be free of hallucinations. There would be many observations and facts concerning the patient which went "into" the diagnosis and yet which, in the form of a diagnosis, would not be available as a basis for evaluating his status when discharged. Likewise, there could be many observations concerning the patient at the time of discharge which would be relevant to predictions of his adjustment at home and at work and yet which would not be properly weighed and communicated by recording only that his hallucinatory behavior had ceased. *Evaluations of this gross type permit recognition of the most obvious degrees of improvement, i.e., abatement of severe symptomatology, but they are not sufficient for detection of improvement in the patient's personality with respect to those inadequacies in social adjustments and diurnal habits which served as the first clues to his illness.* From this point of view, the maximal efficacy of a given therapy for a given patient could not be ascertained at

the time of his discharge from the hospital but only after he had returned to a normal, non-institutional environment and only in terms of the quality of his adjustments in that environment. Furthermore, the observations and evaluations of his post-therapy behavior would have to be ordered to the same method of recording as was used to depict his status prior to therapy. Limitations of time and money have generally prohibited the attainment of this ideal paradigm for the validation of a therapeutic procedure. The next best approach demands that the status of the individual be thoroughly measured in terms of such observations as are possible immediately prior to the start of a course of therapy and immediately upon the completion of that course. It is equally important in this more circumscribed evaluation of the patient's progress that his status before and after therapy be measured with respect to the same variables.

It may be argued that no matter how rigorous the procedure of investigation, the "pure" effects of shock therapy, for example, may not be determined since such therapy is always imbedded in a certain setting and administered by personnel whose effects on the patients cannot be held constant. It may further be argued that the effects assignable to shock are inextricably mixed with the effects of the less dramatic psychotherapeutic efforts of the physician. With regard to this last point, it must be recognized that the ancillary therapeutic programs developed for the patient receiving convulsive therapy are probably as diverse as the theories proposed to explain the mechanics (or dynamics) of shock as therapy. This source of discrepancy should be remembered when it is necessary to explain different measured effects in different clinics when the

same (?) therapy has been applied.

With regard to the complicating effects of other psychotherapeutic efforts (e.g., interviews) accompanying a major medical therapy, it should be recognized that the clinic in which the decision to apply a convulsive therapy is *not* accompanied by a reduction (either planned or unplanned) in the number of hours spent by the physician in personal contacts with the patient is probably rare. When a gross procedure of the nature of electroshock seems a promising approach with a given patient, the therapist probably feels some justification in reducing his contacts with that patient so as to afford more hours to those patients for whom no "specific" therapy is available. In short, the effects of a therapy such as electroshock are probably open to evaluation in a more pure form than would at first appear to be the case.

The design and carrying through of investigations of the above nature must necessarily await the development of adequate techniques for the thorough descriptive assay of the individual personality. It has been only within the last decade that clinical psychology has offered an instrument having the essential virtues of psychiatrically oriented variables empirically validated and quantitatively expressed. This instrument is the Minnesota Multiphasic Personality Inventory (8, 9).² This psychiatric inventory was designed primarily to furnish an external, objective source of information bearing on the detection and diagnosis of personality aberrations. In addition to four scales which indicate the validity of the subject's record, the MMPI measures personality deviations on nine clinical scales: Hs—hypochondriasis, D—depres-

² The Minnesota Multiphasic Personality Inventory will be referred to hereafter as the MMPI.

sion, Hy—hysteria, Pd—psychopathic deviate, Mf—masculinity-femininity, Pa—paranoia, Pt—psychasthenia, Sc—schizophrenia, and Ma—hypomania.

The validity measures are the "Cannot Say," L, F, and K scales. The first of these is simply the number of items placed in the "Cannot Say" category. Since the "Cannot Say" responses are not scored on any of the other scales, a large number of such responses has a generally suppressing effect on the clinical profile. The L scale is composed of fifteen items designed to measure the strength of the subject's tendency to "fake good" by selecting responses which place him in a socially acceptable light. The F scale is composed of items the scored responses to which are extremely infrequent; elevated F-scores may occur as a result of errors in scoring, carelessness or lack of comprehension by the subject, or a "fake bad" tendency in the subject. The K scale is the most recent of the validity scales and provides a measure of the degree to which "defensiveness" has char-

acterized the subject's test-taking attitude.

The wide use of the MMPI by psychiatrists, clinics, and general physicians attests to its success in a diagnostic role (2, 15). It is interesting to note that only recently has there been a sign of appreciation of its potential value in measuring the effects of a given course of therapy (7, 15, 16). That there should be such a lag in the use of a basically diagnostic instrument for the evaluation of change reflects a general failure on the part of psychiatrists and psychologists to recognize that the determination of the patient's condition at any time after hospitalization or the onset of therapeutic interviews cannot be usefully made except on the very same variables the assessment of which served as the basis for hospitalization or the onset of therapy. The clinician should be no less willing to accept supportive information for his post-therapy evaluation than he was to seek assistance in the original problem of diagnosis.

CHAPTER II

PURPOSE, SUBJECTS AND PROCEDURES

THE MAJOR ASSUMPTION underlying the investigation reported here is that an instrument of demonstrated validity in the measurement of personality deviations of psychiatric import is of equal validity for the evaluation of personality changes and of the therapeutic operations deemed responsible for those changes. It is not assumed, however, that a direct evaluation of change in personality, or of the efficacy of a therapy, is possible through the single study of score changes. Rather, it is believed necessary to determine whether there are patterns of change in scores on a multi-variable instrument which are characteristic for given patients under given regimens and to discriminate, on as objective a basis as possible, between those changes in the personality assay which are therapy-related and those which are merely chance. Further, it is believed necessary to evaluate the extent of the relationship between changes in item response and associated score changes.

In the event that an individual completes a personality questionnaire such as the MMPI on two different occasions, certain analyses may be made of any differences which are observed in the resulting profiles. It may be asked if the two profiles differ in a degree which would appear to approach significance. More specifically, this question relates to whether the two scores obtained on each single scale of the profiles differ by such an amount as exceeds that for which the hypothesis of no true difference (between the means of hypothetical populations of first-profile and second-profile

scores) is tenable. For simplicity in reference, this is the problem of *status-score change*. The observed differences between the various pairs of status-scores constitute *change-scores*.

There is still another kind of change not yet extensively considered, perhaps because attention to status-score changes has tended to mask the operation of this more subtle, but potentially significant variation from profile to profile. This is the change in response to specific items composing the various scales; it may be referred to as *item-response change*.

The purpose of this investigation has been to study the item-responses on the MMPI of certain groups of persons and the changes in the item-responses of these persons as a result of the passage of time plus the occurrence of certain events. More specifically, it has been designed to study the nature of changes in the responses to the MMPI items of certain groups of psychiatric patients and the relationships between item-response changes and status-score changes in these groups, and it has been oriented toward the problem of detecting and measuring therapy-induced changes in personality.

The investigation began with a careful check of the approximately 7000 records in the MMPI research files at the University of Minnesota to obtain the names of all persons, normals and patients, to whom the inventory had been administered on at least two occasions. For all cases, available information on education, occupation, marital status, place of residence, etc. was recorded in addition to the interval, to the nearest 5-day unit,

between the two administrations of the MMPI. No in-patients were included in the study unless a check of their hospital record revealed that they had been discharged "improved" or "recovered" and that the two MMPI profiles had been obtained within the limits of a single hospitalization, and the length of this hospitalization was recorded in addition to the above information.¹ The hospital charts of the patients were also checked carefully to determine the exact nature of the therapy which the patient had received, and particular note was made of the use of electroshock.

A post-admission MMPI is obtained routinely on all patients admitted to the Psychopathic Unit of the University of Minnesota Hospitals. With the beginning of this study in September, 1946, a special effort was made to collect post-therapy, pre-discharge MMPI records on all patients. Most of the patients comprising the "psychotic-shock" group for this study were obtained in this manner.

The Psychopathic Unit has a total of only 35 beds of which six are in a separate ward for grossly disturbed patients. While the number of beds available to men and women patients on the general psychiatric ward are approximately equal, the actual number of women hospitalized per year greatly exceeds the number of men.² Due to lack of an adequate number of male patients for whom two records are available, the present study deals entirely with females.

The normal cases fall into two groups: the first is composed of nineteen female night-school students, half of whom took

the individual form of the MMPI first and one week later took the group form, the other half taking the two forms in reversed order with a one-week interval between them.³ The second group is composed of 23 females selected from the files because they had two MMPI records. The mean interval between the two tests for this second group was 136.4 days; the median interval was 47.5 days. This indicates the marked positive skewness which characterized the distribution of inter-test intervals. The mean age of the first group was 32.1 years, with a standard deviation of 7.3; the mean age of the second group was 26.3 years, with a standard deviation of 4.4. The difference in the mean ages of the two groups is not significant ($t = .998$) nor is the difference in the variances significant ($F = 2.727$). The MMPI test and retest means and standard deviations for the two normal groups are given in Table 1. Only one of the test-retest differences in mean score is statistically significant; this difference occurs on the Pt scale in the night-school sample ($t = 3.35$; $P < .01$). The test-retest correlation for the Pt scores in this group is .783, the highest of such correlations in this group for any of the clinical scales. If we may assume, on the basis of the very marked tendency in both groups of normals for the second test means on the clinical scales to be lower than the original means and on the basis of similar findings in other studies which have indicated a tendency for the "adjustment scores" of normals to change in the direction of improvement, that there exists an

¹In the case of the out-patients included in this study, the number of visits to the Out-patient Psychiatric Clinic between the two MMPIs was recorded. No evaluation of "improvement" in these patients was available.

²In 1947, the Psychopathic Unit treated a total of 180 female and 123 male in-patients.

³These cases were originally included in a study to determine the equality of the group and individual forms. No significant differences were found between the mean scale scores obtained with the two forms. The writer is indebted to Dr. Howard P. Longstaff for the data from these cases.

TABLE 1
TEST AND RETEST MEANS AND STANDARD DEVIATIONS OF THE MMPI SCORES OF THE
TWO GROUPS OF NORMALS*

Group	Test	?		L		F		K		Hs		D	
		M	σ	M	σ	M	σ	M	σ	M	σ	M	σ
Night-school	1	12.6	9.7	3.2	2.2	3.4	3.1	55.2	8.6	48.0	5.4	51.4	8.8
	2	17.3	19.0	3.4	2.2	2.6	3.2	56.3	8.6	46.0	5.7	49.0	7.3
File-random	1	20.4	17.8	3.4	2.3	3.0	2.0	57.9	7.8	46.3	7.8	47.0	10.0
	2	21.9	20.4	3.6	2.0	2.4	1.8	59.0	7.8	45.4	5.7	45.3	7.6
		Hy		Pd		Pa		Pt		Sc		Ma	
		M	σ	M	σ	M	σ	M	σ	M	σ	M	σ
Night-school	1	52.1	6.8	54.0	8.8	53.8	10.2	52.1	6.6	53.1	7.5	51.4	8.2
	2	50.9	7.8	52.9	9.2	50.6	9.3	48.7	6.8	51.5	8.4	49.0	8.7
File-random	1	50.4	8.6	49.6	7.1	50.2	8.6	47.8	7.2	52.2	7.8	48.4	9.4
	2	48.2	8.4	49.2	8.5	51.0	7.0	47.2	7.8	50.7	6.2	49.2	10.8

* The MMPI data in this and all subsequent tables are based on T-scores derived from the usual formula:

$$T = 50 + 10 \frac{X - M}{\sigma}$$

except for the data for the "?," L, and F scales which are in raw score units.

as yet unanalyzed factor making for a regression toward the norm in the personality measure of normals, even in the absence of originally extreme scores, then the appearance of such a factor would be maximized in the test-retest differences of those scales having the greatest reliability. This explanation of the change on the Pt scale shown by the

night-school students is offered in the absence of evidence of any more obvious factor which might account for the observed difference. It will be noted from Table 2, which presents a summary of descriptive statistics for the normal and clinical groups, that the combined normal group is characterized by a high incidence of individuals with high school

TABLE 2
GENERAL DESCRIPTIVE STATISTICS FOR THE FIVE SAMPLES TESTED

Group ^a	Number	Age		Days in Hospital ^b		Days between MMPIs		% Single	% H.S. ^c
		Mean	Range	Mean	Range	Mean	Range		
Normals	42	28.9	21-46			61.6	2-1060	78.6	95.0
OPD	25	32.1	17-56	5.8	2-18	80.6	5-330	32.0	63.2
PN	24	26.9	17-39	52.0	20-114	33.4	5-75	41.7	34.8
PS-NS	13	35.8	20-57	64.4	15-106	33.8	5-65	38.5	30.8
PS-S	20	43.2	20-59	50.6	23-90	27.2	5-40	30.8	83.3

^a OPD—25 out-patients.

PN—24 hospitalized neurotics.

PS-NS—13 hospitalized psychotics who received neither insulin nor electroshock therapy.

PS-S—20 hospitalized psychotics treated with electroshock.

^b The figures in this column for the OPD group refer to number of clinic visits in the interval between the two MMPIs.

^c Percent of group with high school education or better.

education or better, from metropolitan areas.

The clinical groups are four in number: 25 out-patients; 24 hospitalized neurotics; 13 hospitalized psychotics who did not receive any type of convulsive therapy; and 20 psychotics who received elec-

TABLE 3
THE COMPOSITION OF THE CLINICAL GROUPS
BY DIAGNOSIS

Group	Diagnosis	Number
OPD N = 25	Psychoneurosis (mostly mixed) . . .	22
	Schizophrenia, simple	1
	Hypomania	1
	Undiagnosed	1
PN N = 24	Psychoneurosis, mixed	13
	hysteria	4
	hypochondriasis	2
	psychasthenia	1
	reactive depression	4
PS-NS N = 13	Manic-depressive, depressed . . .	3
	manic	2
	circular	1
	Agitated depression	1
	Schizophrenia, simple	1
	catatonic	1
	Recurrent depression	1
PS-S N = 20	Psychosis, undiagnosed	3
	Manic-depressive, depressed . . .	4
	Involutional, melancholia	8
	paranoia	2
	Schizophrenia, paranoid	2
	hebephrenia	1
	undetermined	1
	Paranoid condition	2

troshock therapy. Table 3 presents the composition of the clinical groups by diagnosis.

The two MMPI records of each subject were tallied on a common form in such a way as to indicate the actual response ("True," "False," "Cannot Say") of the subject to each item on each of the two testings. From these combined records, it was possible to determine for each individual the total number of changed responses, the number of changes by direction of change, and the number of changes per general type of item. These

data for the individuals were then combined to permit determination of group statistics.

The mean score of each group on each of the validity and clinical scales of the MMPI was determined for the first and second tests and the reliability of the differences between the test and retest means of each group was determined by statistical analysis.

The inter-group differences in the means and variances for the number of item-responses changed per individual were tested for reliability as were the inter-group differences in means and variances for the number of responses changed per item.

In order to study the relationship of item-response changes to changes in status-scores, the total change-score, i.e., change in status-scores, for each individual was determined by adding, without regard to sign, the differences between his test and retest scores on the clinical variables. The ratio of this total change-score (Σcs) to the number of items changed (Σic) was determined for each individual and summary statistics for this ratio ($\Sigma cs / \Sigma ic$) were determined for each of the groups.

Finally, item-by-item comparisons were made to determine the reliability of the differences between the per cent of responses changed by the normal group and each of the clinical groups respectively. Lists of items with reliable differences in frequency of change were compiled and studied to determine if they showed any dominant characteristics.

In view of the small N for each of the patient groups, the present study can only be considered as exploratory rather than definitive, a search for support or refutation of specific hypotheses, rather than a basis for final establishment of conclusions.

CHAPTER III

RESULTS

A. MEAN PROFILES FOR THE NORMAL AND CLINICAL GROUPS

THE MEANS and standard deviations of each of the groups on each of the validity and clinical scales of the MMPI, for both administrations of the inventory, are given in Table 4. The corresponding profiles are presented in Figures 1-5.

The significance of the difference between the first and second administration means for a given scale in a given group was tested by the following formula:

$$t_{(M_1-M_2)} = \frac{M_1 - M_2}{\sqrt{\frac{N \Sigma D^2 - (\Sigma D)^2}{N^2(N-1)}}$$

where D = the difference between the first and second scores of a given individual. Only those differences have been considered reliable which yield a " t " score with a probability of less than 1 per cent.

It will be noted with respect to the clinical variables that there is a general tendency for the mean scores on the second administration to be lower than those for the first. Exceptions to this occur on the Pd scale in the out-patient (OPD) group, on the Mf scale in the hospitalized neurotic (PN) group, and on the D and Hy scales in the psychotic patients who did not receive electroshock therapy (PS-NS). None of the mean score differences of the combined Normal Group is statistically significant nor is any of the OPD group. The mean F -score difference in the Normal Group has a P of $< .02$.

In the PN group, the second test means on the Hs , Hy , and Pd scales are significantly lower than the original means ($P < .01$ in each case).

The PS-NS group shows no significant mean differences on any of the clinical scales. The difference between the two K -score means is significant at the 1 per cent level.

The sample of psychotic patients who received electroshock therapy (PS-S) shows a reliable decrease of mean score on the Hs , D , Hy , Pd , Pt , and Sc scales. The mean difference on the Pa scales has a probability of less than 2 per cent.

With respect to the validity scales, it should be noted that there is a consistent tendency for the F -score mean to be smaller on the second test, while the reverse holds for the K -score except for the PS-S group. None of the mean differences for these two variables is reliable, however.

From Table 2, it may be seen that the Normal Group is characterized by a higher proportion of single persons and persons with a high school education or better than characterizes the clinical groups. From Fig. 1 and Table 4, it appears that these selective factors have not unduly influenced the "normality" of this group insofar as the clinical scales of the MMPI are concerned. The two profiles follow the 50-score line closely and there is not the elevation on the Hy -scale to be expected in a college group. Furthermore, the test-retest reliabilities for the various scales in the combined normal group closely approximate those reported for normals in the literature.

In working with personality scores in

TABLE 4
TEST-RETEST MEANS AND STANDARD DEVIATIONS FOR THE SCORES OF EACH GROUP ON THE
MMPI, WITH SIGNIFICANCE TESTS FOR THE MEAN DIFFERENCES^a

Group	Test	?		L		F		K		Hs		D	
		M	σ	M	σ	M	σ	M	σ	M	σ	M	σ
Normals N=42	1	17.0	15.2	3.2	2.2	3.2	2.6	56.6	8.3	47.1	6.0	49.0	9.8
	2	19.8	19.9	3.4	2.1	2.6	2.6	57.8	8.3	45.7	5.7	47.0	7.7
OPD N=25	1	14.8	10.0	4.4	2.2	7.0	4.0	51.3	9.0	74.4	13.4	73.5	13.4
	2	8.0	8.0	4.2	1.6	6.2	3.4	52.4	8.2	70.1	12.6	72.7	9.0
PN N=24	1	14.2	18.6	4.6	2.6	6.1	3.9	52.2	9.8	74.9	13.4	71.8	8.3
	2	11.6	12.4	4.2	2.4	5.2	5.0	53.4	8.5	66.5	12.4	67.6	9.7
PS-NS N=13	1	27.4	23.4	4.3	2.2	10.0	9.4	44.6	6.0	58.6	13.3	64.0	11.0
	2	11.4	9.3	4.5	1.3	6.6	6.7	52.5	6.8	57.8	11.2	66.7	7.3
PS-S N=20	1	14.4	13.6	4.6	2.4	8.2	6.0	52.1	7.2	62.8	10.7	74.4	10.8
	2	23.0	23.4	3.7	2.0	4.0	2.5	50.7	6.9	53.0	12.8	58.2	9.1
		"f"		"f"		"f"		"f"		"f"		"f"	
Normals		1.4		.8		2.1		1.0		1.6		2.1	
OPD		2.0		.8		1.6		.7		2.0		.3	
PN		1.4		1.0		.5		.6		3.2*		1.4	
PS-NS		2.4		1.6		1.3		3.0*		.2		.5	
PS-S		1.7		1.8		3.2*		1.0		4.0*		4.2*	

The data for the "?", L, and F scales are in raw score units. Those for the remaining scales are in T-score units.

^a The N's on which the means and standard deviations for the Mf score are based are: OPD = 20; PN = 13; PS-S = 20.

* Significant at the 1% level.

which deviation is permitted in one direction only, in which experience with like measures indicates a general phenomenon of regression in the direction of "normality" (i.e., lower scores) with a simple passage of time, and in which the expected direction of change for patients receiving specialized therapy is toward lower scores on a retest, we would expect the retest scores of patients to show a reduced range and to find this reduced range reflected in lower test-retest correlations than are found with normals tested over a similar interval. In general these expectations are borne out by the results in Table 5, which show a tendency for the test-retest correlations to be lower for the patient groups than for

the normals. In view of this argument, it may be considered more proper to consider test-retest correlations in patients who have had therapy as indices, at least in part, of therapeutic efficacy rather than as reliability coefficients. The small size of the samples in this study precludes testing the significance of the differences between these coefficients but their values are reflected in the tests of the reliability of the differences in corresponding means since they enter into the determination of the standard error of the difference between means for the test and retest scores. In this respect, it should be noted that Darley (4) interpreted his test-retest coefficients, which were considerably lower than the corrected

TABLE 5
TEST-RETEST CORRELATIONS OF EACH GROUP ON EACH SCALE

Group	?	L	F	K	Hs	D	Hy	Pd	Mf	Pa	Pt	Sc	Ma
Normals (N=42)	.718	.810	.749	.663	.724	.734	.682	.569		.585	.701	.594	.783
OPD (N=25)	.430	.556	.797	.580	.657	.424	.664	.568	.827	.367	.798	.796	.746
PN (N=24)	.884	.596	.687	.763	.536	.570	.617	.605	.580	.498	.475	.608	.819
PS-NS* (N=13)	.334	.852	.375	.105	.220	.707	.436	.375		.052	.147	-.393	.608
PS-S (N=20)	.434	.555	.384	.353	.908	.486	.555	.431	.343	.580	.496	.660	.450

* The coefficients for the PS-NS group are product-moment "r's" corresponding to computed "rho's" as given in Table 45, p. 362, Garrett, *Statistics in Psychology and Education*, 2nd edition.

TABLE 4—Continued

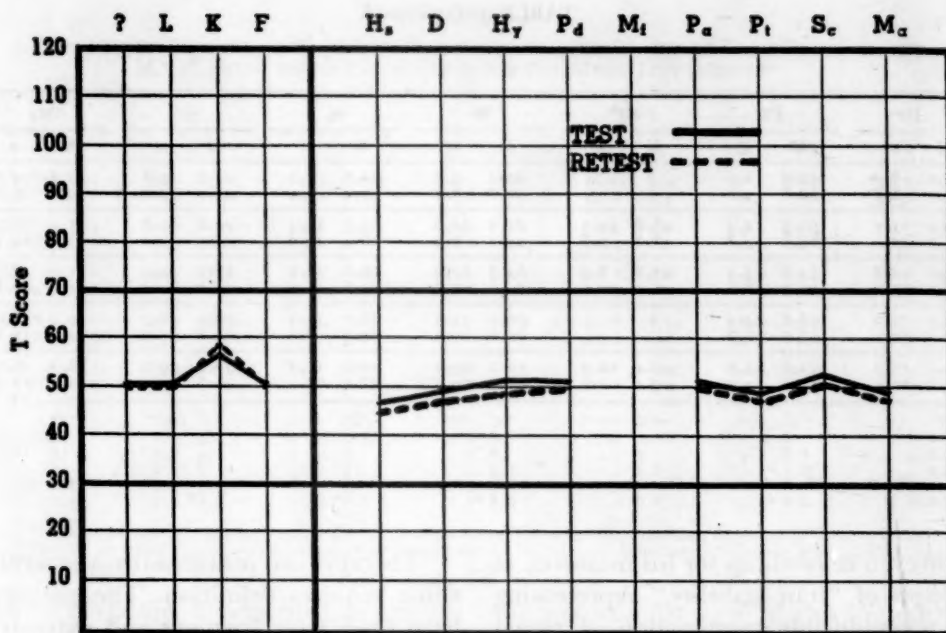
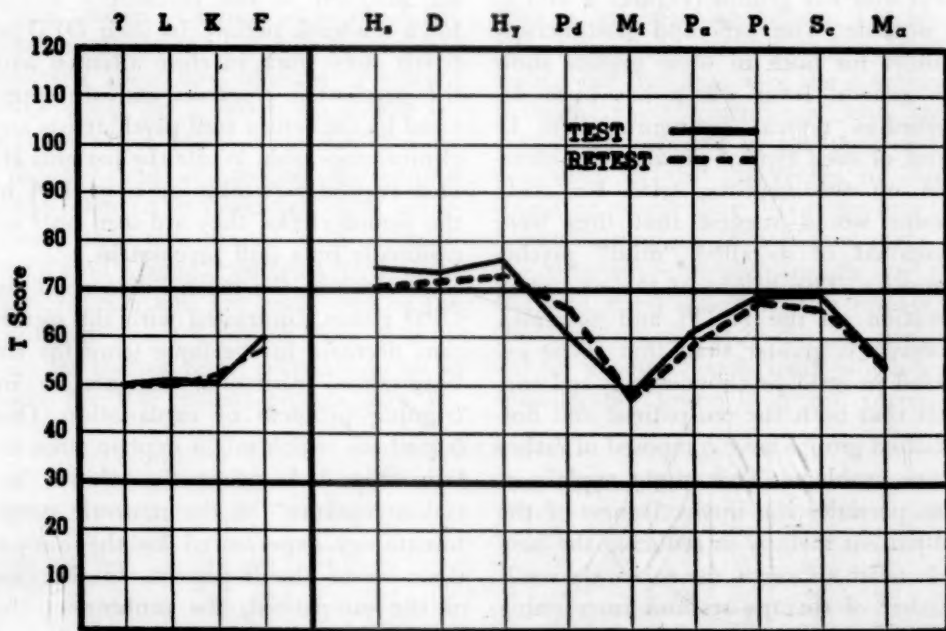
Hy		Pd		Mf ^b		Pa		Pt		Sc		Ma	
M	σ	M	σ	M	σ	M	σ	M	σ	M	σ	M	σ
51.6	8.0	51.6	8.2			51.8	9.3	49.8	7.2	52.6	7.6	49.8	9.0
49.4	8.2	50.8	9.0			50.8	8.2	47.9	7.4	51.0	7.2	49.1	9.8
76.2	11.2	61.3	12.2	48.8	12.1	61.2	10.6	69.5	11.4	69.8	12.8	55.9	11.9
73.1	12.4	66.2	14.8	48.8	9.8	59.7	9.4	68.4	13.6	67.3	13.2	55.0	10.5
77.2	10.8	63.6	13.2	42.8	8.9	62.5	10.6	66.6	11.8	69.2	12.4	55.2	12.7
69.2	11.3	60.5	11.0	46.2	10.6	59.1	10.2	65.8	12.7	67.0	14.4	55.8	12.6
62.0	12.0	58.8	14.4			70.0	12.3	66.2	11.4	70.0	16.3	59.0	18.0
63.1	10.4	58.1	6.8			66.0	14.4	63.4	8.6	64.8	15.6	53.3	17.8
65.2	13.5	72.0	15.8	49.4	10.0	70.2	14.2	70.9	17.2	72.2	19.7	56.2	12.2
55.7	12.2	57.6	13.3	49.4	9.2	59.2	19.4	57.6	15.3	59.8	15.7	52.5	12.3
"f"		"f"		"f"		"f"		"f"		"f"		"f"	
2.0	.6	.6				.8		2.0		1.4		.6	
1.6	1.8	1.8		.1		.8		.5		1.4		.6	
4.0	4.2	4.2		1.4		1.5		.2		.8		.4	
.2	.2	.2				.6		.8		.7		1.2	
3.4	4.2	4.2		1.6		2.6		3.5*		3.6*		1.2	

odd-even reliabilities for his measures, as gauges of "trait stability," representing a non-reducible combination of actual change in his subjects plus errors of measurement.

The similarity of the profiles for the OPD and PN groups (Figures 2 and 3) is notable. The pre- and post-therapy profiles for both of these groups show the general form which has been described as "typical" for neurotics (6). In terms of such typical profiles, the elevation on the neurotic triad for these groups would suggest that they were composed of so-called "mild" psychoneurotics. The height of the secondary elevation on the Pa, Pt and Sc scales, however, is greater than that to be expected in mild psychoneurotics and suggests that both the out-patient and hospitalized groups were composed of rather severe problems. This might explain at least partially the ineffectiveness of the out-patient therapy in reducing the neurotic triad although the relatively small number of therapy sessions intervening between the two profiles (see Table 2) is undoubtedly a factor also.

"Therapy" as provided by the OPD clinic requires definition. The patients have their most frequent and extensive contacts (for history-taking, therapeutic interviews, etc.) with "junior clerks." These are junior medical students who are assigned to the psychiatric service for a ten-week period. In their OPD activity, they work in close alliance with the psychiatric residents and are supervised by the senior staff psychiatrists and clinical associates. While the patients are seen regularly (weekly or bi-weekly) by the junior clerks, they are seen only occasionally by a staff psychiatrist.

The rise in the mean Pd score of the OPD group, contrasted with the significant decrease in the same score for the hospitalized neurotics, presents an intriguing problem of explanation. One hypothesis which might explain these relationships is based on the relative "social acceptance" of the neurotic symptomatology experienced by the out-patient versus the in-patient. In the case of the out-patient, the probing of the therapist to uncover possible conflictual material which might have an etiologic

FIGURE 1. Mean test and retest profiles of the combined normals. ($N = 42$.)FIGURE 2. Mean pre-therapy and post-therapy profiles of the out-patients. ($N = 25$.)

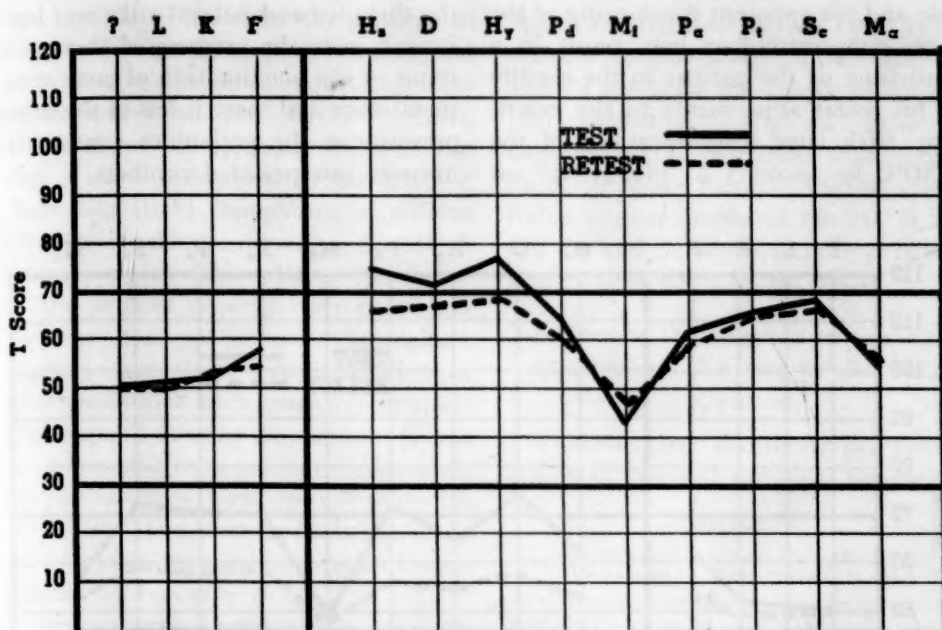


FIGURE 3. Mean pre-therapy and post-therapy profiles of the hospitalized neurotics. (N = 24.)

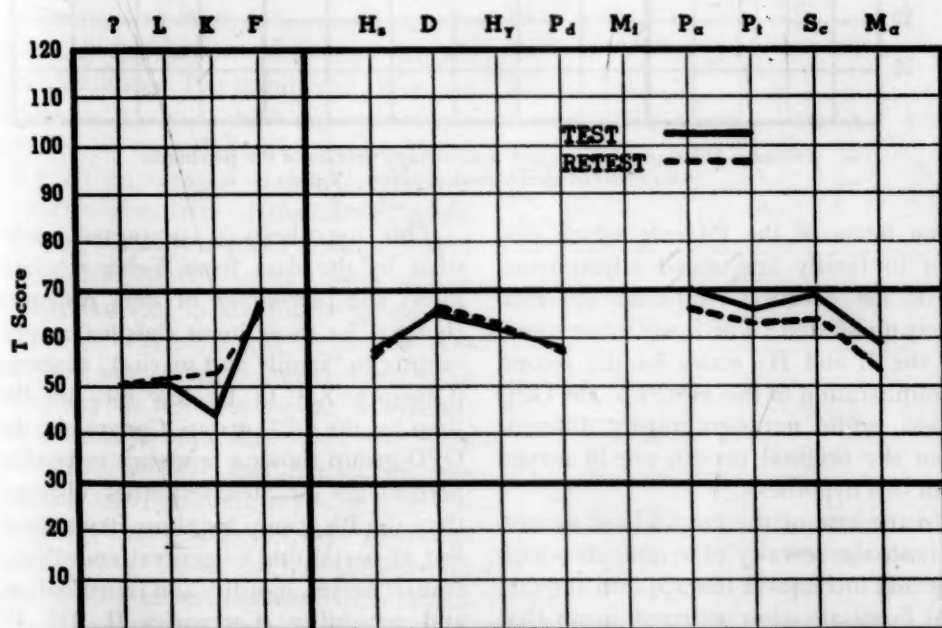


FIGURE 4. Mean pre-therapy and post-therapy profiles of the psychotics who did not receive electroshock therapy. (N = 13.)

role and his apparent discounting of the gross symptomatology may result in a sensitizing of the patient to the quality of his social adjustments to the extent that, with later administrations of the MMPI, he becomes a "plus-getter" on

the therapist and patient alike may legitimately note the progress of therapy in terms of the amelioration of gross symptomatology and there is less of the subtle pressure on the patient to continue to uncover interpersonal conflicts.

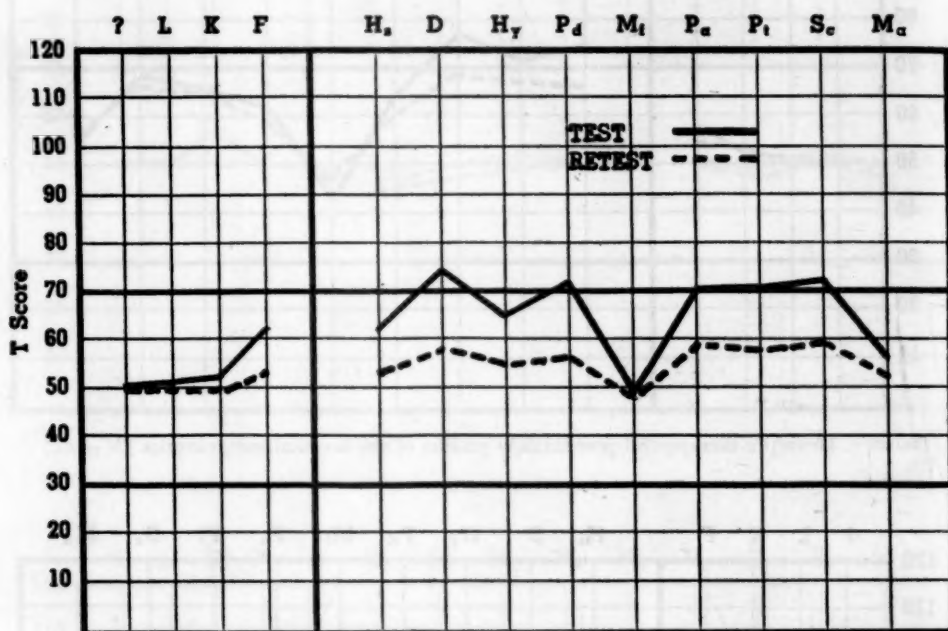


FIGURE 5. Mean pre-therapy and post-therapy profiles of the psychotics who received electroshock therapy. ($N = 20$.)

those items of the Pd-scale which pertain to family and social adjustments, while his neurotic symptoms are relatively unaffected.¹ The lower mean scores on the L and Hy scales for the second administration of the MMPI to the OPD group, while not significantly different from the original means, are in accord with this hypothesis.

In the case of the hospitalized neurotic both the severity of symptoms which justified intramural therapy and the fact that hospitalization occurred, mean that

¹ The term "plus-getter" has been applied to those individuals with a specific motivation, conscious or otherwise, to put themselves in an unfavorable or socially undesirable light.

This hypothesis is supported somewhat by the data from Table 9 which shows the percentage of item responses changed for those items classified as pertaining to "family and marital" concerns (Category XI) to be less for the PN than for the OPD group. Conversely, the OPD group shows a tendency to smaller percentages of item-responses changed than the PN group for those items classified as pertaining to general neurologic, cranial nerves, motility and coordination, and sensibility (Categories II, III, IV and V).

It is interesting to note that with psychotic patients, the ordinary regimen of

the psychiatric ward in the absence of electroconvulsive therapy is not adequate to achieve a significant reduction of the mean scores on the psychotic scales. While all of the psychotic scales in Figure 4 show reductions in mean score, there is a slight though not significant increase in the mean scores for the D and Hy scales. The relative "mildness" of the profiles plus the fact that these patients did not receive electroshock suggests that they were borderline or mild psychotics with extensive neurotic coloring and this fact may account for the apparent shift from psychotic to neurotic symptomatology. Concomitant with this shift, there is an increase in defensiveness or control as revealed in a mean K-score increase of 7.90 ($t = 3.000$, $P < .01$).

Figure 5, which presents the two mean profiles for the group of psychotic patients who received electroconvulsive therapy is of interest not only because of the marked changes it reveals but because of the stability shown by the Mf and Ma scales. The stability of these two scales in a group which shows so much change on the other clinical variables is of great importance as evidence for the selective sensitivity of the MMPI in reflecting therapy-induced changes. Had all of the MMPI scales shown significantly reduced means following electroshock therapy, it might be argued that all that was being measured was a gross tendency to response-change as a result of the confusion and memory defects associated with EST and that specific personality changes were masked. The differential change reflected in the profiles of Figure 5 argues for the efficiency of the MMPI as an instrument for detecting the particular results of special therapies.

In summary, the data on the mean pre- and post-therapy profiles reveals no

significant changes for the OPD group, reliable improvement in the Hs, Hy, and Pd scores for the sample of hospitalized neurotics (PN), no significant changes for the psychotics who received no electroshock therapy (PS-NS), and reliable improvements on the Hs, D, Hy, Pd, Pt, and Sc scores for the psychotics treated with electroshock (PS-S).

B. GROUP DIFFERENCES IN THE FREQUENCY AND NATURE OF RESPONSE CHANGES

Table 6 gives the mean per cent of responses changed per item for each of

TABLE 6
MEANS AND STANDARD DEVIATIONS FOR PER CENT OF RESPONSES CHANGED

Group	N	M	σ	
Normals	42	14.0	10.0	
OPD	25	22.0	10.8	
PN	24	18.4	13.6	
PS-NS	13	30.6	13.2	
PS-S	20	30.6	13.4	

<i>F</i> -values				
Group	OPD	PN	PS-NS	PS-S
Normals	1.1	1.8*	1.7*	1.8*
OPD		1.6*	1.5*	1.6*
PN			1.0	1.0
PS-NS				1.0

<i>t</i> -values				
Group	OPD	PN	PS-NS	PS-S
Normals	12.0*			
OPD				
PN			14.4*	14.0*
PS-NS				.1

* Significant at 1% level.

the groups. These values are the same as the per cent of all item-responses per group which are changed.² The per-

² Let i = no. of items, p = no. of persons in group, and c = no. of responses changed per item. Then the mean values of Table 6 are derived from

$$\frac{\sum(c/p)}{i}, \text{ which is also equal to } \frac{\sum c}{pi}$$

centage of responses changed by the normals agrees well with the lower limits of comparable values reported in the literature. The percentage of responses changed in the two subgroups of normals were very similar (14.41 per cent and 15.56 per cent). It should be noted that the OPD and PN groups not only show an increase in the per cent of responses

and PS-S groups are significantly greater than that for the PN group. The difference between the means of the two psychotic groups is not reliable. An hypothesis suggested by the data of Table 6 is that neurotic and psychotic patients have specific frequencies of response change to the MMPI items which are significantly different from one another, signifi-

TABLE 7
PERCENTAGE DISTRIBUTIONS OF ITEM RESPONSE CHANGES IN EACH GROUP ACCORDING TO DIRECTION OF CHANGE

Group	T to F	F to T	T to ?	? to T	F to ?	? to F
Normal	37.3	33.8	6.2	5.2	9.8	7.7
OPD	42.2	43.7	2.4	4.2	2.2	5.3
PN	46.5	40.4	1.7	3.1	3.2	5.1
PS-NS	45.9	36.2	1.8	5.5	1.5	9.1
PS-S	40.6	40.2	5.4	3.2	6.1	4.5
Group		T to F and F to T		T or F to ?		? to T or F
Normal		71.1		16.0		12.9
OPD		85.9		4.6		9.5
PN		86.9		4.9		8.2
PS-NS		82.1		3.3		14.6
PS-S		80.8		11.5		7.7

changed over that for the normals but that this per cent is practically the same for both of the therapy groups, the difference between the two values being only 3.5 per cent. Likewise, the per cent of responses changed is nearly identical for the two psychotic groups and is twice that for normals, both the shock and non-shock treated psychotics changing approximately one-third of their responses. The mean per cent of responses changed per item for the OPD group is significantly greater than that for the Normal Group. The significantly greater variances of the PN, PS-NS, and PS-S groups as compared to that of the normals prevent testing of the significance of the differences among the related means. The mean values for the PS-NS

cantly greater than that for normals, and relatively independent of the specific type of therapy they are receiving.

Table 7 presents data on the frequency of response change in each of the groups studied in terms of the direction of the changed response. Two facts are immediately apparent from this table. First of all, there appear to be no gross differences among the groups in the directional distribution of the response changes. For each group, the percentage of responses changed from "True" to "False" is within ten points of the percentage of "False" to "True" responses. Secondly, the percentage of response changes which involve directly opposite responses, i.e., from "True to False" or "False to True," constitute, in all groups,

70 per cent or better of all changes. It may be noted that changes involving a shift from an originally uncertain attitude or introspection (as evidenced by a "Cannot Say" response) to a definitely positive or negative ("True" or "False") response constitute from 7.7 per cent to 12.9 per cent of the response changes in the five groups. Since the number of "True" and "False" responses on the first administrations of the MMPI constitute, on the average, 97 per cent of all responses, with the remaining 3 per cent of responses in the "Cannot Say" category, it is clear from Table 7 that original "True" and "False" responses contribute a smaller proportion of the total of changed responses than would be expected from the proportion they constitute of all original responses. On the other hand, the per cents of all response changes which are in the "Cannot Say" to "True" or "False" direction are from two to three times the proportions which the "Cannot Say" responses are of all original responses. From this data, it appears that the "Cannot Say" responses, which are not scored on any of the MMPI scales, are considerably less stable than the "scorable" "True" and "False" responses. This finding is in agreement with that of Lentz (13) who found the proportion of "?" answers changed on the Bernreuter inventory to be approximately twice the proportions of "Yes" and "No" answers which were changed. These findings indicate that any investigation of the differentiating potentiality of "Cannot Say" responses relative to certain classes of psychiatric patients would most likely show such responses to reflect variable rather than stable aspects of the personality. The average proportion of changed responses which are in the "Cannot Say" to "True" or "False" direction is 10.0 per cent for the

four clinical groups: this value may be compared with one of 12.9 per cent for the normals. There is no indication that therapy-induced changes in personality are reflected in an increased tendency for originally uncertain attitudes to become either positive or negative.

The means and standard deviations for the number of item-responses changed per individual for each of the

TABLE 8
MEANS AND STANDARD DEVIATIONS FOR NUMBER OF ITEM-RESPONSE CHANGES PER INDIVIDUAL

Group	N	M	σ
Normals	42	74.4	23.6
OPD	25	108.6	33.0
PN	24	106.5	29.4
PS-NS	13	151.6	61.6
PS-S	20	151.6	37.3

F-values				
Group	OPD	PN	PS-NS	PS-S
Normals	2.0	1.6	7.2*	2.6*
OPD		1.2	3.6*	1.2
PN			4.6*	1.6
PS-NS				2.8

t-values				
Group	OPD	PN	PS-NS	PS-S
Normal	19.2*	18.6*		
OPD		.8		13.3*
PN				10.2*

* Significant at 1% level.

groups is given in Table 8. The means for the two subgroups of normals were 71.37 and 77.04 respectively, with standard deviations of 19.36 and 26.24. The differences between the means and variances of these two subgroups is not significant ($t = 2.424$; $F = 1.818$). Again, as was noted for the over-all per cent of responses changed, there is a remarkable similarity of the OPD and PN groups and of the PS-NS and PS-S groups with respect to the mean number of responses

changed per individual. The mean for the Normal Group is significantly less than that of the OPD group and also significantly smaller than that of the PN group. The variances of the PS-NS and PS-S groups differ significantly from that

the general categories of items, as originally classified by the MMPI authors, was determined for each of the groups. These data are given in Table 9, together with the rank of each category for frequency of response change within

TABLE 9
PERCENT OF ITEM-RESPONSES CHANGED PER GENERAL CATEGORY OF ITEM FOR EACH GROUP*

Category	Normals	OPD	PN	PS-NS	PS-S
I	15.3 (9)	24.0 (7.5)	21.8 (14)	31.6 (10)	38.3 (2)
II	8.0 (22)	23.4 (13)	24.3 (7)	30.8 (12.3)	30.8 (10)
III	16.0 (8)	16.7 (22)	18.9 (18)	34.3 (4)	25.0 (18.3)
IV	7.5 (23)	30.7 (1)	31.9 (2)	37.2 (1)	25.8 (16.5)
V	7.1 (24)	26.4 (3)	35.8 (1)	29.2 (18.5)	32.0 (8)
VI	10.7 (17)	19.2 (17)	22.5 (12)	26.2 (20)	27.0 (14)
VII	10.5 (19)	16.8 (20.5)	23.3 (9.5)	29.2 (18.5)	19.0 (23)
VIII	6.5 (25)	24.0 (7.5)	27.7 (3)	23.8 (21)	25.0 (18.3)
IX	11.9 (13.5)	29.6 (2)	19.2 (17)	32.3 (8.5)	21.0 (22)
X	13.5 (12)	19.4 (16)	21.9 (13)	34.0 (5)	30.3 (12)
XI	10.6 (18)	20.0 (15)	15.4 (22)	29.3 (17)	23.7 (21)
XII	22.4 (1)	22.4 (14)	27.1 (4)	36.3 (2)	39.2 (1)
XIII	11.5 (16)	19.0 (18)	14.9 (24)	23.1 (22)	25.0 (18.3)
XIV	18.3 (5)	23.8 (10)	26.0 (5)	29.8 (15)	31.3 (9)
XV	18.0 (6)	16.4 (23)	17.3 (19)	31.2 (11)	25.8 (16.5)
XVI	21.0 (2)	23.6 (11)	21.6 (15)	32.9 (7)	30.7 (11)
XVII	17.6 (7)	23.5 (12)	23.3 (9.5)	30.8 (12.3)	34.5 (5)
XVIII	11.9 (13.5)	25.3 (4)	23.0 (11)	30.8 (12.3)	34.7 (4)
XIX	18.9 (4)	24.8 (6)	23.8 (8)	34.6 (3)	34.0 (7)
XX	14.1 (11)	17.6 (19)	16.1 (20)	32.3 (8.5)	29.0 (13)
XXI	9.4 (21)	16.1 (24)	15.9 (21)	29.5 (16)	26.1 (15)
XXII	14.5 (10)	23.9 (9)	21.0 (16)	30.2 (14)	34.3 (6)
XXIII	11.6 (15)	12.0 (25)	10.1 (25)	19.8 (23)	18.6 (24)
XXIV	20.7 (3)	25.0 (5)	25.3 (6)	33.3 (6)	35.3 (3)
XXV	9.5 (20)	16.8 (20.5)	15.0 (23)	16.4 (24)	23.7 (20)

* The figures in parentheses give the ranks of the categories in the order of frequency of changed responses within each clinical group.

of the normals and hence testing of the related mean differences is not applicable. The mean for the OPD group does not differ significantly from that of the PN group, and the means of the PS-NS and PS-S groups do not differ significantly. The PS-S group mean is significantly larger than that of either the OPD or the PN group.

In an attempt to determine whether there were any group differences in the pattern of response-changes in terms of the general types of items, the percentage of item-responses changed in each of

each of the groups. The general categories plus the number of items classified in each are as follows:

- I. General Health (9 items).
- II. General Neurologic (19 items).
- III. Cranial Nerves (11 items).
- IV. Motility and Coordination (6 items).
- V. Sensibility (5 items).
- VI. Vasomotor, Trophic, Speech, Secretory (10 items).
- VII. Cardiorespiratory (5 items).
- VIII. Gastrointestinal (11 items).
- IX. Genitourinary (6 items).
- X. Habits (20 items).
- XI. Family and Marital (29 items).
- XII. Occupational (18 items).
- XIII. Educational (12 items).

- XIV. Sexual Attitudes (19 items).
XV. Religious Attitudes (20 items).
XVI. Political Attitudes—Law and Order (46 items).
XVII. Social Attitudes (72 items).
XVIII. Affect, Depressive (32 items).
XIX. Affect, Manic (24 items).
XX. Obsessive, Compulsive (15 items).
XXI. Delusions, Hallucinations, Illusions, Ideas of Reference (31 items).
XXII. Phobias (29 items).
XXIII. Sadistic, Masochistic (7 items).
XXIV. Morale (33 items).
XXV. Items modeled after suggestions of Hartshorne, May and Shuttleworth to indicate whether the individual is trying to place himself in an improbably acceptable or unacceptable light (15 items).

Application of the χ^2 -technique revealed no significant difference in the distributions of frequencies of response changes per general category for the two subgroups of normals ($\chi^2 = 27.34$; $P > .30$).

While the most "unstable" categories in the normal subjects are XII (Occupational) and XXIV (Morale), those with greatest frequency of response change in the OPD and PN groups are IV (Motility and Coordination) and V (Sensibility). Greatest frequency of change for the PS-NS and PS groups is in XII (Occupational). There is no clear differential pattern of frequency of response change per category of item. In general, the first nine categories, those dealing with physical symptomatology of the various organ systems, Category XVIII (Affect, Depressive) XXI (Delusions, Hallucinations, Illusions, etc.) XXII (Phobias), and XXV (the "L" scale items) show a considerably greater frequency of response change in the clinical groups than characterizes the normals.

C. ITEM CHANGE IN RELATION TO SCORE CHANGE

In order to make an appraisal of the relationship between the amount of item-response change and the degree of

profile change, the profiles of the three individuals having the largest number of item-response changes in each group were studied together with the profiles of the three individuals with the smallest numbers of response changes. It was at once apparent that no close relationship existed between the number of item responses changed and the degree of profile change.

This brief examination suggested that there were individual differences and probably group differences in the relative specificity of item changes with respect to profile changes which might be interpreted as reflecting therapeutic progress. It was thought advisable to devise a measure of the relationship of item-change to scale-score change and to study the distribution of this measure in each of the special groups. The measure used for this purpose was simply the ratio of the total number of item-responses changed per individual to the total amount of change, without regard to sign, in the raw scores on the clinical variables. Table 10 reports the means and standard deviations of this ratio for the normals and for each of the clinical groups.

Since only 213 of the 495 items of the MMPI included in this study are scored on any of the eight psychiatric variables covered by those items (Hs, D, Hy, Pd, Pt, Sc, and Ma), only 43 per cent of item changes would be expected to be reflected in scale score change. Thus, for every 10 item-responses changed, a minimum change of 4.3 score points would be expected on the clinical scales, if the changes in item-response were random or chance. If, on the other hand, the item-changes were in response to particular therapeutic efforts directed at the deviations measured by certain of the psychia-

tric scales, the ratio of item-change to score-change should be lessened, i.e., a given number of item-changes should result in a greater amount of scale-score change than is true for normals.

The findings are reported in Table 10.

The Normal Group shows a mean ratio of item-change to score-change

TABLE 10
MEANS AND STANDARD DEVIATIONS FOR RATIO
OF NUMBER OF ITEMS CHANGED TO AMOUNT
OF SCORE CHANGE

Group	N	M	σ
Normals	42	4.0	2.5
OPD	25	3.3	1.0
PN	24	3.2	1.2
PS-NS	13	3.4	1.4
PS-S	20	3.4	1.7

F-values				
Group	OPD	PN	PS-NS	PS-S
Normals	5.8*	3.8*	3.0	2.1
OPD		1.5	1.9	2.6
PN			1.2	1.8
PS-NS				1.4

t-values				
Group	OPD	PN	PS-NS	PS-S
Normals			2.6*	3.6*
OPD		1.3	.3	.4
PN			1.1	.9
PS-NS				.0

* Significant at the 1% level.

which is just about what is to be expected by chance. The means for the clinical groups are smaller, indicating an increasing specificity of item-change, an increasingly higher proportion of scale-score items among those which are changed. The mean ratios for the PS-NS and PS-S groups are significantly smaller than that for the Normal Group. The difference between the mean ratios of the OPD and PN groups is not significant nor is that between the means of the PS-NS and PS-S groups.

One sees, then, that neurotic and psychotic patients undergoing various therapies not only tend to change a significantly greater number of their item-responses than do normals (Table 8), but that their response changes show a significantly greater concentration of items which are scored on the psychiatric scales of the MMPI. We have in this data positive evidence of the potential validity of the MMPI for the measurement of therapeutic effects. At the same time, the lack of close relationship between over-all amount of item change and profile change argues for caution in the interpretation of profile changes. For example, if an alcoholic takes the MMPI when not fully recovered from the general confusional and somatic effects of a severe drinking bout and repeats the inventory later with a more clear sensorium, one would expect a very large number of response changes and probably considerable profile change. No matter how extensive the profile changes might be under these conditions, they should not be interpreted as evidence of a change in personality but rather as evidence of recovery from an acute disturbance of certain perceptual-cognitive and somatic functions. Thus, if the F-score of the original profile were very high, it need not necessarily invalidate the profile in the usual sense, but rather might better be interpreted as indicating that the clinical profile reflects an acute disturbance and does not typify the individual's basic personality. If, on the other hand, this same person were tested and retested when fully sober, over an interval in which intensive psychotherapy had been attempted, an equally extensive change in his MMPI profile might be found to result from a change in response to considerably fewer items than were changed in the above example.

True *personality* changes as reflected in profile changes should mean a greater-than-chance frequency of change of those items which yield scale scores.

D. CHARACTERISTICS OF THE MOST STABLE ITEMS AND THE MOST VARIABLE ITEMS IN EACH GROUP

Since the variability of the items (percentage of responses per item changed) shows a large range in each of the groups

found on the F and Sc scales. A similar relationship, concentration of the most stable items on the F and Sc scales, holds in each of the four clinical groups.

No equally consistent pattern is found among the groups for the most variable items. In the normals and PS-NS groups, the Hy scale has the highest concentration of variable items. In the PN group, the Hy and Sc scales have the highest—and equal—concentrations of variable

TABLE 11
PERCENTAGE DISTRIBUTIONS, BY SCALE, OF THE SCORES YIELDED BY THE MOST STABLE AND BY THE MOST VARIABLE ITEMS FOR EACH GROUP

Group	Item Type	No. of Items	% Change	Tot. Scale Score	Percent of Total Scale-Score Per Scale											
					L	F	K	Hs	D	Hy	Pd	Mf	Pa	Pt	Sc	Ma
Normal	Stable	27	0	38	5.3	34.2		5.3	5.3	7.9	2.6	2.6	10.5	5.3	18.4	2.6
	Variable	28	35.0+	25		4.0	20.0		8.0	24.0	12.0	4.0	12.0	4.0		12.0
OPD	Stable	30	0-4	39	5.1	48.7		2.6	2.6	2.6	7.7		12.8		17.9	
	Variable	37	40.0+	31			16.1		9.7	9.7	16.1	9.7	12.9	9.7		16.1
PN	Stable	49	0-4.2	55	3.6	47.3		1.8	1.8	1.8	5.5	1.8	10.9	1.8	23.6	
	Variable	33	41.7+	45	2.3		4.4	11.1	11.1	17.8	8.9	4.4	6.7	11.1	17.8	4.4
PS-NS	Stable	28	0-7.7	31	9.7	22.6		6.5	9.7	9.7	3.2	12.9		6.5	16.0	3.2
	Variable	37	53.8+	35			14.3		8.6	17.0	14.3	2.9	14.3	8.6	5.7	14.3
PS-S	Stable	20	0-5	30	6.7	53.3					6.7		6.7	3.3	23.3	
	Variable	32	55.0+	37			5.4	8.1	21.6	16.3	16.3		5.4	10.7	8.1	8.1

(e.g., 0-42.99 per cent in the normals), it was considered desirable to determine what differences might characterize the most stable and the most variable items for each distribution. Table 11 presents the results of an analysis of the scale representations of the two types of item in each group. Thus, it was found that the 27 items which had a 0 per cent change in the Normal Group (i.e., each individual's second response to an item was the same as his first response) yielded a total of 38 raw scale-score units.³ Of the total amount of scale-score yielded by these items, the largest percentages are

items. In the PS-S group, the D scale shows the highest concentration of variable items, with the Hy and Pd scales showing the second highest frequency. In short, the F and Sc scales are marked in all groups by a relatively high frequency of very stable items; the variable items, on the other hand, are irregular in their scale concentration from group to group. In view of the stability characterizing many of the items composing the F and Sc scales, one would have reason to expect any change in the scores on those scales to more probably reflect valid personality changes rather than to be chance variations.

Table 12 reports for each group the proportions of the most stable and most variable items for that group which are "obvious," "subtle," and not scored on

³ That the number of scale-score units should exceed the number of items is accounted for by the fact that certain of the items are scored on more than one scale. Lists of the most stable items, i.e., those with "0%" change, are provided in the Appendix.

any scale. The "obvious" and "subtle" items are composed of all items scored on the D, Hy, Pd, Pa, and Ma scales of the MMPI and have been keyed by Wiener and Harmon (18) according to the relative ease or difficulty with which the nature of the socially desirable response may be detected. It will be noted that there is a consistent tendency for the

non-scored items in the stable and variable items is significant for the Normal and OPD groups. In short, the variable items, as distinguished from the stable items, are characterized by a tendency to be either of a type in which the direction of the socially desirable response is not readily perceptible or of a type which has not yielded normal-criterion

TABLE 12
THE PROPORTIONS OF THE MOST STABLE AND MOST VARIABLE ITEMS FROM EACH GROUP WHICH ARE "OBVIOUS," "SUBTLE," AND NOT SCORED ON ANY SCALE

Group	The Most Stable Items				The Most Variable Items			
	No. of Items	Percent Obvious	Percent Subtle	% Not Scored	No. of Items	Percent Obvious	Percent Subtle	% Not Scored
Normal	27	25.9	0.0	14.8	28	10.7	32.1	47.9
OPD	30	23.3	6.7	13.3	37	27.0	16.2	48.6
PN	49	12.2	6.1	20.4	33	27.3	21.2	21.2
PS-NS	28	21.4	3.6	28.5	37	21.6	32.4	30.0
PS-S	20	10.0	0.0	20.0	32	40.6	15.6	40.6

stable items to consist of higher proportions of obvious rather than subtle items, with a mean difference of 15.3 per cent between the proportions of these two types of item among the stable items of the five groups. No such consistent tendency appears in the obvious-subtle distributions of the variable items. In comparing the stable and variable items, no consistent group-to-group relationship appears with respect to the frequency of obvious items. However, in all groups, the variable items show higher proportions of subtle items than are found in the stable items, and for the Normal and PS-NS groups the differences between the proportions of subtle items in the stable and variable items respectively are significant ($P < .01$). Furthermore, there is a consistent tendency for the variable items to have a higher proportion of items which are not scored on any scale than is true for the stable items, and the differences between the proportions of

critical ratios sufficiently large to result in their selection for scoring on any of the existing MMPI scales. That items having a high frequency of changed responses upon repetition with a given group should show these characteristics is to be expected both from the psychology of that behavior which is variable and from the nature of item-selection. Thus, there are areas of behavior the "subtlety" of which is expressed by the absence of any clear-cut, readily recognized pattern of response in the group which may serve as a constant standard against which the individual can evaluate or regulate his own conduct. In the absence of such social norms, the individual is required to "project" norms which vary as a function of recent experience and memory of last response in a similar situation. The element of uniqueness in all experience plus the fallibility of memory makes directly for variability of response in the individual.

This variability expresses itself in the group by an absence of a given response in sufficient frequency to characterize that group in statistical comparison with another group.

As an aside, it might be proposed that the ideal questionnaire should be com-

ing effects to operate in such a way that they may be measured, rather than naively assuming that they have been circumvented by a clever ruse.

Further study of the stable and variable items was made in an effort to discover any distinguishing characteristics

TABLE 13

THE PROPORTIONS OF ITEMS INVOLVING FREQUENCY ESTIMATES AND ITEMS NOT INVOLVING SUCH ESTIMATES IN THE STABLE AND VARIABLE ITEMS OF EACH GROUP

Group	Type of Item	No. of Items	% With Freq. Estimate	% Without Freq. Estimate	Ratio of No. of Non-Freq. to Freq. Items
Normal	Stable	27	23.1	76.9	3.33
	Variable	28	33.3	66.7	2.00
OPD	Stable	30	16.7	82.3	5.00
	Variable	37	40.5	59.5	1.47
PN	Stable	49	14.3	85.7	6.00
	Variable	33	38.2	61.8	1.62
PS-NS	Stable	28	31.0	69.0	2.33
	Variable	37	37.8	62.2	1.64
PS-S	Stable	20	10.0	90.0	9.00
	Variable	32	28.1	71.9	2.55

posed solely of items of the subtle variety which require the subject to "project" his own norms, thus gaining for the personality inventory the highly regarded virtues of the projective techniques. There are at least two answers to this proposal. In the first place, the variability in response to such items, as has been indicated in this study, would suggest that they might seriously detract from the effort to establish reliable relationships between the test measures and given criteria. Secondly, recognizing that there are attitudes peculiar to the process of test-taking which are different from the particular attributes we are interested in measuring and which have certain distorting effects on the latter, there would seem to be some inherently scientific cogency to permitting such distort-

that might not have appeared in the above statistical analyses. Other studies (10) have indicated that items involving an estimate of frequency by the subject (e.g., "much of the time," "seldom," etc.) have higher rates of change than items not requiring such estimates. An analysis was made of the occurrence of these two types of item among the stable and variable items of each of the groups in the present investigations, with the results reported in Table 13. A consistent tendency was found for the stable items of each group to have a higher proportion of items not involving an estimate of frequency than was true for the variable items. For the stable items from all groups, the mean ratio of the number of items not involving an estimate of frequency to the number of items involving

such an estimate was 5.13; for the variable items, the comparable mean ratio was 1.86. It appears from this that there is a distinct tendency for those items in which the subject must make some estimate of frequency of occurrence of a behavior in himself or others (e.g., "I feel weak all over much of the time," or "I believe I am no more nervous than most others") to be more variable in the responses they elicit from the same individuals than are the items not requiring such estimates.

In a study of the relationship of variability of items and the degree to which they were stated in present or past tense, no consistent group-to-group tendency was found. In three of the five groups (Normal, PN, and PS-S), the variable items yielded a higher ratio of "number of items in the present tense" to "number of items in the past tense" than did the stable items. In the other two groups (OPD and PS-NS), this relationship was reversed, the stable items yielding a higher ratio of present-tense to past-tense items. The mean of these ratios in the five groups was 4.10 for the stable items and 8.10 for the variable items. As might be expected, those items involving a memory of past occurrences appear more frequently among those items to which responses are quite stable, while those items involving present situations, moods, and pressures appear more frequently among those items to which responses are changed. However, the variability in the relative concentration of present-tense and past-tense items among the stable and variable items from group to group suggests that the relationship between the "time" factor of an item and its stability or variability is not a simple one but involves other elements.

Finally, a study was made to determine the relative stability of responses to

the so-called "double-negative" items. These are items which are stated in the negative and are scored if answered in the negative, that is sorted "False" (e.g., "I do not tire quickly"). The stable and variable items from each of the five groups of subjects were examined to determine the number of double-negative items they included. This survey of the 154 stable and 157 variable items yielded a total of 33 items of the double-negative variety. Of these 33 items, 10 (or 30.3 per cent) were stable and the remainder, 69.7 per cent, were variable. In four of the five groups, the ratio of the number of double-negative items among the variable items to the number among the stable items was three to one. In the remaining group (PS-NS), the ratio was one to one.

In summary, the above analyses of the most stable and most variable items for each of the five groups of subjects have pointed to certain fairly consistent differentiating characteristics. The stable items tend to be concentrated on the F and Sc scales of the MMPI, to be obvious in nature, to involve no estimate of the frequency of occurrence of a given behavior either in the subject or in other people, and to be more frequently stated in past rather than present tense. The variable items, on the other hand, tend to show no clear concentrations on any of the MMPI scales, to be subtle rather than obvious, to involve estimates of frequency of given behavior, and to be stated in the present rather than the past tense.

E. INVESTIGATION OF A THERAPY-SUSCEPTIBLE SCALE

It was suggested earlier in this research that one possible direction to be taken by a study of item-response changes would be a determination of those items which

have a significantly greater percentage of changed responses in persons responding favorably to therapy than in normals. Theoretically, such a scale of items would have both a prognostic and an evaluative function. For example, the greater the clinical score made by the individual on those items having a specific tendency to change with electroshock therapy, the greater the probability that his post-therapy responses will be in the non-psychiatric direction, i.e., the greater the probability that he will improve. It follows directly from this that the amount of the individual's change in score on the therapy-susceptible items may be taken as an index of the validity of the overall profile change as an indication of improvement. Profile changes in the favorable direction not accompanied by adequate changes on the items specifically responsive to the therapy concerned would not be misinterpreted as evidence of improvement.

It might be argued that such proposed scales of therapy-susceptible items should be composed solely of those items for which the frequency of change of an originally significant response in a given therapy group was reliably greater than the same change in normals. Such a selection procedure would miss those equally variable items which happened on the first test to have been answered in the non-significant direction. Consequently, the following procedure was used to arrive at a list of "therapy-susceptible" items for each of the treatment groups.

The total number of response-changes was determined for each item for each of the "improved" clinical groups (PN, PS-NS, PS-S) and for the combined normals.⁴ These frequencies were converted

to percentages. By the use of appropriate nomographs (19) the reliability of the difference between the percentage of changed responses for each item in the Normal Group and in a given therapy group was determined. Only those differences were considered reliable which yielded a critical ratio of 3.00 ($P < .003$). In the cases of percentage differences which were borderline with respect to the nomographs, the following formula was applied:

$$t = \frac{p_1 - p_2}{\sqrt{\frac{N_1 p_1 + N_2 p_2}{N_1 + N_2} \left(1 - \frac{N_1 p_1 + N_2 p_2}{N_1 + N_2} \right) \left(\frac{1}{N_1} + \frac{1}{N_2} \right)}}$$

By this procedure, three scales of "therapy-susceptible" items were derived. In terms of the arguments given above, the items of these scales may be considered to constitute a potential measure of prognosis and the scales will be referred to hereafter as the Px scales.

Since the hospitalized neurotics appeared to be the group for which independent validating data could be most readily obtained, it was decided to study the prognostic effectiveness of the scale of items derived for this group.⁵ A key based on these items was applied to both MMPI records of the 24 patients in the PN group. Basic statistics for the Px scores in this group are given in Table

which item changes were distributed over the twenty-five general categories of items.

A study of the significance of the difference of frequency of response-change per item in the two normal groups revealed a total of only 15 items which yielded differences significant at the 5 per cent level. Of these 15 items, only 5 (1 per cent of the total number of items) yielded differences significant at the 1 per cent level. In view of these findings, it was considered permissible to combine data from the two normal groups for the item study.

⁵The twenty-four items comprising this Px scale, as numbered in the MMPI manual of directions, 1943 version, are as follows: A2, A4, A13, A16, A46, A48, B1, B9, B11, B12, B14, B15, B17, B28, C43, F2, F18, F49, F55, G1, G20, G21, H46, I27.

⁴It has been reported previously that the χ^2 technique showed no significant differences between the two normal groups in the manner in

14. The mean Px score on the second MMPI is significantly lower than that obtained in the first administration. While this is evidence of the potential usefulness of the Px scores for detecting therapeutic changes on a group basis, the extreme overlap of scores before and

discharge in either the "improved" or "unimproved" groups ($\chi^2 = .939$ and $.254$, respectively).

In order to determine whether any better differentiation would result from the use for the Px scale of only those items for which the frequency of change of an originally significant response was reliably greater in the PN group than in the Normal Group, a key was prepared for the 11 items from the 24 of the first Px scale which satisfied this relationship. These items were: A-2, A-13, A-46, B-9, B-11, B-14, B-17, B-28, F-18, and I-27. The results of the application of this scale to the two test groups of hospitalized neurotics revealed no improvement in differentiation over that of the original Px score ($t = .446$; $F = 1.076$).

It would appear from the above data that the general approach to the selection of therapy-susceptible items which has been followed in this study does not result in a valid set of items having any prognostic power to differentiate pa-

after therapy makes it practically useless for individual prediction.

In order to study the distribution of the Px score in independent samples, the hospital files were searched for cases of diagnosed female psychoneurotics who were discharged "unimproved." Only thirty such cases were found and of this number, only nineteen had MMPI records. Twenty-nine cases of female psychoneurotics discharged "improved" and having post-admission MMPI records were also chosen randomly from the files. Table 15 gives the basic statistics of the Px score for these two groups.

The difference between the mean Px scores of the two groups is not significant.

In addition to the information on "condition on discharge" for the two test groups, the diagnostic summary of each patient carried a statement relative to prognosis. These statements were evaluated in terms of "good," "fair," or "poor" prognosis. The χ^2 technique revealed no relationship between the Px score at the time of admission and the prognosis on

tients who will improve under a given regimen from those who will not. This conclusion necessarily assumes that the criterion used to test the validity of such items was in itself a valid criterion.

In the nine years that the Psychopathic Unit of the University Hospitals has been in operation, a total of 470 female admissions has been diagnosed as

TABLE 14
MEANS, STANDARD DEVIATIONS, AND
SIGNIFICANCE TESTS FOR THE Px SCORE
IN THE HOSPITALIZED NEUROTICS
(PN) ($N = 24$)

	M	σ	Range
1st MMPI	11.9	3.5	4-20
2nd MMPI	8.1	4.8	1-21
$t = 3.5$ $P < .01$			
$F = 2.1$ $P < .05 > .02$			

TABLE 15
MEANS, STANDARD DEVIATIONS, AND
SIGNIFICANCE TESTS FOR THE Px SCORE
IN SAMPLES OF IMPROVED AND
UNIMPROVED PSYCHONEUROTICS

	M	σ	Range
Improved ($N = 29$) . .	10.6	4.6	3-20
Unimproved ($N = 19$)	8.6	5.1	0-18
$t = 1.4$ $P < .20 > .10$			
$F = 1.2$ $P > .05$			

psychoneuroses. Of this total, only 6.4 per cent were discharged "unimproved." A recovery or improvement rate of better than 90 per cent is so much greater than the median value reported in the literature that one would necessarily suspect a biased selection of patients, a therapeutic program of particular composition and intensity, an inaccurate evaluation of post-therapy status, or an admixture of these factors.

Careful evaluation of the histories of the neurotic patients in this study and of the larger hospital sample of neurotics from which they were drawn does not support the hypothesis of biased selection of case material. Review of the procedures used in the treatment of patients by the Psychopathic Unit and analysis of the hospital regimen, aside from particular therapeutic techniques, does not support the hypothesis that the greater improvement rate in the patients of this study as contrasted with those reported in the literature is to be accounted for by differences in treatment. There remains the probability that the high frequency

of improved patients chiefly reflects errors of omission in the clinical evaluation of the status of patients and the amount of their change at the time of discharge. Support to this point is given by a readmission rate of 10.5 per cent for the female neurotics.

A more objective determination of the post-therapy status of patients than is usually available is greatly to be desired. Such determination ultimately must reside in the clinical judgment of the psychiatrist. To the extent that the psychiatrist can be encouraged to orient his post-therapy evaluation of a patient to a relationship with the factors underlying his assay of pre-therapy status, he will provide a valid clinical criterion. With such a criterion available, it will be possible for research techniques to establish independent prognostic and evaluative measures which will not only have value in the consideration of the individual patient but will provide even more important objective checks on the efficacy of specified therapeutic efforts.

CHAPTER IV

SUMMARY AND CONCLUSIONS

THIS STUDY has had three goals: to determine the nature and significance of changes in the MMPI profiles of groups of psychiatric patients receiving certain therapies, to study the extent and nature of item-response changes in these clinical groups and to contrast them with those found in normals, and, finally, to attempt the selection and testing of a group of therapy-susceptible items.

A check of the approximately 7000 records in the MMPI research files at the University of Minnesota yielded a total of 124 cases for which two profiles were available and which met the other criteria for this study. These cases were distributed as follows: normals, 42; out-patients, 25; hospitalized neurotics, 24; hospitalized psychotics not treated with any type of shock therapy, 13; hospitalized psychotics given electroshock therapy, 20. All in-patients included in the sample had been discharged "improved" or "recovered" and the two MMPI profiles had been obtained within the limits of a single hospitalization. Since it was not possible to find a sufficient number of male cases, the samples were composed entirely of females.

The sample of psychotic patients treated with electroshock, though a very mixed group, was composed largely of depressions and involutional melancholias. The average number of grand mal treatments received by these patients was between six and seven, and the interval between their two MMPI's was slightly less than one month. This group showed a significant post-shock reduction of their mean scores on all of the MMPI clinical scales with the exception

of Masculinity-Femininity and Hypomania. Those means which were highest before therapy in this group were on the Depression, Psychopathic Deviate, Psychasthenia and Schizophrenia scales, and it was on these scales that greatest change occurred.

The sample of hospitalized neurotics had a mean initial profile practically identical with that of the out-patient sample. However, after a mean interval of approximately one month, the same as the median interval for the out-patient group, the profile for the hospitalized neurotics showed a significant decrease of the Hypochondriasis, Hysteria and Psychopathic Deviate scores. It must be remembered that all patients in the three hospitalized groups were discharged as "improved" or "recovered." It may be concluded that improvement in the personality pattern of neurotics under the general regimen of the psychiatric ward is shown in a significant reduction of hypochondriacal and hysterical symptomatology as measured by the MMPI. Apparently, this form of therapy does not suffice to reduce any marked degree of depression as reflected in the Multiphasic "D" score. The ineffectiveness of hospitalization alone to cure depressive states in a neurotic syndrome can only be judged with respect to the interval of intramural therapy represented in our sample, namely, approximately one month.

Within the limits of our small sample of psychiatric out-patients and the nature and number of psychiatric interviews intervening between the collection of the two MMPI records, it would not appear

that out-patient therapy of the type afforded by the University Clinic suffices to affect a significant change in the mean scores of such patients on any of the clinical scales. The out-patient sample consisted almost wholly of psychoneurotics of mixed type who had an average of five clinic visits between their tests. No clinical evaluation of the nature of their response to therapy was available. The data of this study would indicate that such patients require more intensive therapy than that afforded by the out-patient clinic or a greater number of therapy interviews (or both) before they will show a favorable response. This conclusion is presented with certain reservations. It may be argued that improvement in the out-patient group occurred in areas of attitude, self-evaluation, and personal relationship not sampled by the MMPI. It may further be argued that the validity of the MMPI as a measure of therapeutically derived changes in personality has not been sufficiently established to warrant its application as the sole criterion in evaluating such changes. However, in the absence of other criteria and with consideration for the validity of the MMPI as a picture of the patient's pre-therapy status, the data of the present study do not yield evidence of improvement in the psychiatric status of the out-patients.

With respect to the proportion of item-responses changed, it was found that all of the clinical groups changed a larger percentage of their responses than was true for the normals, who changed only 14 per cent of their responses. Both groups of neurotics changed approximately 20 per cent of their responses, while the two psychotic groups gave a 30 per cent response change. Aside from the fact that the clinical groups show a

greater variability than the normals in their responses, it is of considerable interest that both groups of neurotics show essentially the same rate of change in spite of marked differences in treatment, and that both psychotic groups likewise show a nearly identical mean frequency of response change in spite of the marked differences in the therapy they received. These findings suggest a characteristic rate of variability for normals, for neurotics, and for psychotics which is relatively independent of environmental stresses, although the items on which the variability is based are different. On a stability-of-response continuum, the neurotics would fall between the normals and the psychotics, somewhat closer to the normals. Further investigation of this finding might yield a useful technique for solving borderline neurotic-psychotic differentiations.

It has been shown that for normals and for groups of neurotics and psychotics, "Cannot Say" responses have a considerably higher rate of change than is found for either "True" or "False" responses. The proportion of all changed responses contributed by changes in original "Cannot Say" responses is from two to three times greater than the proportion which "Cannot Say" responses are of all original responses. The greater variability of the "Cannot Say" responses suggests that such responses may occur on items which pertain to components of personality which are "validly" variable rather than stable.

By selecting cases from the extremes of the distributions of number of items changed per individual, it was demonstrated that no simple relationship holds between the total number of item-responses changed and the amount of scale-score change effected. Marked

changes in profile may be accomplished by individuals who have changed relatively few items, whereas other individuals may change a considerably greater number of items and yet show very little profile change.

This finding led to a study of the ratio of the number of items changed to the total amount of change in scale-score for the members of each of the groups. As a result of this investigation, it was discovered that the clinical groups not only change a significantly greater number of items than do normals, but that the items they change include a significantly greater concentration of scored items, i.e., items contributing to a score on at least one scale of the MMPI. This finding is considered evidence of the sensitivity of the MMPI to therapy-induced changes in personality.

While the most stable items for each of the groups, normal and clinical alike, tended to be concentrated in the F and Schizophrenia scales, no similar pattern of concentration on any scale or scales characterized the most variable items.

With respect to the "obvious" or "subtle" qualities of the items, stable items for each group were much more often obvious than subtle. No similar relationship was found for the most variable items, but these items were much more frequently not scored on any of the MMPI scales than were the stable items. The tendency for the most variable items to be more commonly not scored on any of the scales probably reflects the fact that when item selection is on the basis of group differentiation at a given time, items which pertain to the more subtle and fluctuating aspects of personality, and hence on which there would be no marked agreement by a given group, are missed.

The attempt to derive a scale of items which would both predict the response of a given type of patient to a given therapy and measure the amount of personality change to be attributed to that therapy was not successful. By selecting those items with a frequency of response-change reliably greater in an improved patient-therapy group than in normals and by keying such items for the psychiatrically significant response, it was believed that the resultant scale would yield scores with a positive relationship to prognosis, high scores being associated with good prognosis and low scores with poor. The scale derived by this approach for the hospitalized neurotics was applied to both sets of MMPI records for this group. In accordance with the assumptions about the therapy-measuring potential of such a scale, the mean score of the post-therapy records for the hospitalized neurotics was significantly lower than that of the pre-therapy records, although there was a marked degree of overlap in the two distributions of the score, over a fifth of the post-therapy scores being above the mean of the pre-therapy distribution.

Application of the therapy-susceptible scale for hospitalized neurotics to independent samples of patients discharged "improved" and "unimproved" respectively yielded no differentiation. It was pointed out in this connection that the extreme rarity of hospitalized neurotics who were discharged "unimproved" would give reason to question the validity of the "improved-unimproved" criterion. Within both the "improved" and "unimproved" groups no relationship was found between the score on the therapy-susceptible scale at time of admission and the prognosis at time of discharge.

Several lines of research are suggested by the results of this study. It is obvious, of course, that the analyses reported here should be repeated for data on males. With respect to the detection of non-specific profile changes and their discrimination from therapy-reflecting changes, further study of the characteristics of profiles for people with high and low ratios of number of items changed to amount of score change should be made. The frequency of item-response changes in other samples of neurotics and psychotics should be determined as

a check on the hypothesis that there is a characteristic rate of variability in these groups which might be used as a differential sign.

The problem of developing a prognostic scale for the MMPI would appear to be seriously hampered by the lack of a reliable clinical criterion of patients' responses to therapy. Until a technique may be developed whereby the "condition on discharge" will carry more exact meaning than it appears to currently, scales to predict response to therapy will be very difficult to validate.

APPENDIX

THE MOST STABLE ITEMS IN EACH GROUP¹

A. THE TWENTY-SEVEN ITEMS WITH "0%" CHANGE IN THE COMBINED NORMALS (N = 42):

A11, A19, A40, A46, B5, B32, B46, C13, C14, C36, C49, D29, E37, G11, G35, H4, H14, H15, H20, H21, H22, H25, H30, H48, I7, J50, J53.

B. THE TWELVE ITEMS WITH "0%" CHANGE IN THE OUT-PATIENT GROUP (N = 25):

A51, D43, E21, E22, G44, G53, H14, H15, H20, H27, J45, J48.

C. THE FIFTEEN ITEMS WITH "0%" CHANGE IN THE HOSPITALIZED NEUROTICS (N = 24):

C9, C14, C46, C52, D21, D25, D43, E22, E25, H8, H15, H20, H22, H24, H41.

D. THE SEVEN ITEMS WITH "0%" CHANGE IN THE "NON-SHOCK" PSYCHOTICS (N = 13):

E15, E20, E49, H38, I9, J48, J51, J53.

E. THE SIX ITEMS WITH "0%" CHANGE IN THE "SHOCKED" PSYCHOTICS (N = 20):

C11, C14, C52, D28, H24, I9.

¹ The items are numbered as in the MMPI manual of directions, 1943 version.

BIBLIOGRAPHY

1. APPEL, J. S., BEEBE, G. W., AND HILGER, D. W. Comparative incidence of neuropsychiatric casualties in World War I and World War II. *Amer. J. Psychiat.*, 1946, **103**, 196-199.
2. BENTON, A. L., AND PROBST, K. A. A comparison of psychiatric ratings with MMPI scores. *J. abn. soc. Psychol.*, 1946, **41**, 75-78.
3. CARMICHAEL, H. T., AND MASSERMAN, J. H. Results of treatment in a psychiatric outpatient department—a follow-up study of 166 cases. *J. Amer. med. Assoc.*, 1939, **113**, 2292-2298.
4. DARLEY, J. G. Changes in measured attitudes and adjustments. *J. soc. Psychol.*, 1938, **9**, 189-199.
5. FELIX, R. H. The National Mental Health Act. *Ment. Hyg.*, 1947, **31**, 363.
6. GOUGH, H. G. Diagnostic patterns on the MMPI. *J. clin. Psychol.*, 1946, **2**, 23-27.
7. HALES, W. M., AND SIMON, W. Minnesota Multiphasic Personality patterns before and after insulin shock therapy. *Amer. J. Psychiat.*, 1948, **105**, 254-258.
8. HATHAWAY, S. R., AND MCKINLEY, J. C. A multiphasic personality schedule: I. Construction of the schedule. *J. Psychol.*, 1940, **10**, 249-254.
9. HATHAWAY, S. R., AND MCKINLEY, J. C. *Manual for the Minnesota Multiphasic Personality Inventory*. New York: Psychol. Corp., 1945.
10. HERTZMAN, M., AND GOULD, R. The functional significance of changed responses in a psychoneurotic inventory. *J. abn. soc. Psychol.*, 1939, **34**, 336-350.
11. KELLY, T. L. *Statistical methods*. New York: Macmillan, 1923.
12. KUBIE, L. W. Elements in the medical curriculum which are essential in the training for psychotherapy. *Transactions of the First Conference on Training in Clinical Psychology*, March 27-28, 1947. Josiah Macy, Jr., Foundation.
13. LENTZ, T. F. Reliability of opinionnaire technique studied intensively by the retest method. *J. soc. Psychol.*, 1934, **5**, 338-364.
14. McCANN, W. H. Responses of psychotic patients to the Thurstone Personality Schedule before and after metrazol treatment. *Ind. Acad. Sci.*, 1938, **48**, 187-189.
15. MODLIN, H. C. A study of the MMPI in clinical practice. *Amer. J. Psychiat.*, 1947, **103**, 758-769.
16. PACELLA, B., PIOTROWSKI, Z., AND LEWIS, N. D. C. The effects of electroconvulsive therapy on certain personality traits. *Amer. J. Psychiat.*, 1947, **104**, 83-91.
17. WALKER, H. N. Concerning the standard error of a difference. *J. educ. Psychol.*, 1929, **20**, 53-60.
18. WIENER, D. N., AND HARMON, L. R. I. Subtle and obvious keys for the Minnesota Multiphasic Personality Inventory; their development. *Office Memo #2 to Chiefs of Guidance Centers*. Veterans Administration. Vocational Rehab. & Educ. Div., Minneapolis, Minn. August, 1946.
19. ZUBIN, J. Nomographs for determining the significance of the difference between the frequencies of events in two contrasting series of groups. *J. Amer. stat. Assoc.*, 1939, **34**, 539-544.